ECARES

Revealed Preference and Aggregation

Laurens Cherchye
CES, KULeuven

Ian Crawford
Institute for Fiscal Studies and University of Oxford

Bram De Rock
SBS-EM, ECARES, Université libre de Bruxelles

Frederic Vermeulen
Department of Economics, KULeuven

March 2015

ECARES working paper 2015-08

ECARES
ULB - CP 114/04
50, F.D. Roosevelt Ave., B-1050 Brussels BELGIUM
www.ecares.org

Revealed Preference and Aggregation

Laurens Cherchye[*], Ian Crawford[†], Bram De Rock[‡] and Frederic Vermeulen[§]

March 13, 2015

Abstract

In the tradition of Afriat (1967), Diewert (1973) and Varian (1982), we provide a revealed preference characterisation of the representative consumer. Our results are simple and complement those of Gorman (1953, 1961), Samuelson (1956) and others. They can also be applied to data very readily and without the need for auxiliary parametric or statistical assumptions.

Acknowledgments: We are very grateful to Richard Blundell, Don Brown, Martin Browning, Erwin Diewert, Arthur Lewbel, Peter Neary, Krishna Pendakur and seminar participants at CIREQ Montreal, Keele, Leuven, Oxford and Paris Dauphine for helpful comments and suggestions. Previous versions of this paper have circulated under the titles "Aggregation without the aggravation? Nonparametric analysis of the representative consumer" (CentER Discussion Paper, 2011-143) and "Gorman revisited: nonparametric conditions for exact linear aggregation" (KU Leuven CES Discussion Paper, DPS13.05).

1 Introduction

As noted by Chiappori and Ekeland (2011) "the notion of aggregation is pervasive in economics". It has, of course, long been a core question and, in particular, there has been an important literature on the circumstances under which it is possible to treat aggregate demand as if it were the outcome of choices being made by a single, rational, optimising, normatively

[*]Center for Economic Studies, University of Leuven. E. Sabbelaan 53, B-8500 Kortrijk, Belgium. E-mail: laurens.cherchye@kuleuven.be. Laurens Cherchye gratefully acknowledges the European Research Council (ERC) for his Consolidator Grant 614221. Part of this research is also funded by the FWO (Research Fund-Flanders).

[†]Institute for Fiscal Studies and University of Oxford, Department of Economics, Manor Road Building, Manor Road, Oxford OX1 3UQ, United Kingdom. E-mail: ian.crawford@economics.ox.ac.uk. Ian Crawford gratefully acknowledges financial support from the ESRC Centre for the Microeconomic Analysis of Public Policy (RES-544-28-5001) at the Institute for Fiscal Studies.

[‡]ECARES , Université Libre de Bruxelles. Avenue F. D. Roosevelt 50, CP 114, B-1050 Brussels, Belgium. E-mail: bderock@ulb.ac.be. Bram De Rock gratefully acknowledges the European Research Council (ERC) for his starting grant 263707. Part of this research is also funded by BELSPO.

[§]Department of Economics, University of Leuven, Naamsestraat 69, B-3000 Leuven, Belgium. E-mail: frederic.vermeulen@kuleuven.be. Frederic Vermeulen gratefully acknowledges financial support from the Research Fund KU Leuven through the grant STRT/12/001 and from the FWO through the grant G057314N.

significant, representative consumer. These circumstances are known to be very demanding. The earliest results are due to Antonelli (1886) but the best known in this area are probably those of Samuelson (1956) and of Gorman (1953, 1961), who derived the conditions under which aggregate demand can be written as a function of prices and aggregate income alone. Gorman showed that such exact aggregation is possible if and only if a particular *shape restriction* holds: the Engel curves of consumers are all straight lines with a common slope.[1] Moreover, he showed that exact aggregation implies the existence of a normatively significant representative consumer.

In this paper, we revisit some of the basic questions in the theory of aggregation. However, we do this from a rather different perspective, that of the revealed preference tradition of Samuelson (1938, 1948), Afriat (1967), Diewert (1973) and Varian (1982). Rather than describing the restrictions on choice behaviour in terms of shape restrictions on certain not-directly-observable functions (symmetry of the cross derivatives of the consumer's cost function, or linearity of Engel curves, for example), this approach works by characterising them in terms of a finite system of inequalities involving only the prices and the consumer's observed choices.

The characterisation of important aggregation results in terms of revealed preference inequalities is of theoretical interest, and we present a sequence of closely linked results which provide this, but this is not our only motivation. Our motivation is also empirical. Revealed preference methods directly analyse the raw data themselves. In contrast, methods based on shape restrictions require that the relevant functions are known, and since we never observe functions, these have to be estimated. The conclusions from such an exercise necessarily rest jointly on the validity of the hypothesis at stake plus a number of crucial auxiliary statistical assumptions necessary to deliver consistent estimates of the functions of interest. Revealed preference methods are, to a great extent, free of these auxiliary hypotheses, and so allow researchers to focus with much greater clarity on the economic hypothesis at the core. Furthermore, they are applicable when there are very few observations and hence when statistical methods would be infeasible or too imprecise. As such, these methods can be used by empirical researchers to assess the empirical validity of exact aggregation without unnecessarily aggravating the analysis.

The cost of the revealed preference approach is that, due to its "nonparametric" nature, its empirical restrictions can be relatively weak compared to methods which assume full knowledge of Engel curves, cost functions, and the like. In the present context this might turn out to be an advantage. This is because the microeconometric evidence, based as it is on shape restrictions, has been strongly anti the representative consumer. Papers which consider the question of whether or not the representative consumer exists have therefore tended to be rather negative (see especially Kirman, 1992, and Carroll, 2000). The greater empirical flexibility of the revealed preference approach, by contrast, has the potential to allow us to reassess this result.

This paper is structured as follows. In Section 2 we introduce the notation and some core concepts with respect to the individual consumer. Then, we distinguish between the positive and the normative representative consumer (only the latter plays a meaningful role from a

[1] Exact linear aggregation is to be distinguished from exact nonlinear aggregation, where aggregate demand is a function of some representative level of aggregate income, which itself can be a function of the distribution of income over the individuals (see Muellbauer, 1975, 1976).

welfare economics point of view). We also state the revealed preference conditions for the existence of a normative representative consumer given a socially optimal income distribution rule (following Samuelson, 1956). In addition, we derive the revealed preference conditions for a slightly strengthened definition of a normative representative consumer. This requires a specific assumption on the distribution of either the marginal utility of income at the micro level or the social weights at the macro level. Importantly, because our characterisation in Section 2 is defined for a given income distribution rule, it does not guarantee exact linear aggregation (which requires that aggregate demand depends *only* on the aggregate income). However, it does provide a useful first step towards establishing the revealed preference conditions for such exact aggregation. This is discussed in Section 3, which contains the core contribution of this paper. Specifically, we here investigate the link between the conditions derived in Section 2 and the well-known Gorman aggregation conditions. Along the way we also provide a revealed preference characterisation of Gorman Polar Form preferences for an individual consumer (which, thanks to a knife-edge result, can be surprisingly weak from an empirical point of view) and, based upon this characterisation, we propose an easy-to-implement necessary and sufficient test for the existence of a normative representative consumer that holds for all possible income distributions across consumers. Interestingly, we can show that this test is empirically equivalent to the test developed in Section 2 (for aggregation à la Samuelson, 1956, assuming a socially optimal income distribution) under a fairly weak data requirement. Section 4 offers a summary and some conclusions.

2 Positive and normative representative consumers and income distribution rules

In this section we introduce some first concepts and results that will be useful for our following discussion. We start by briefly reviewing the revealed preference conditions for rational consumption behaviour of individual consumers. Next, we make the distinction between the positive and the normative representative consumer, and we will argue that the latter concept is the only meaningful one from a welfare economics perspective. Subsequently, we derive necessary and sufficient conditions for the existence of such a normative representative consumer for a given, socially optimal income distribution rule. Essentially, this provides a revealed preference treatment of the aggregation concept originally considered by Samuelson (1956). It sets the stage for our discussion in Section 3, where we will consider the revealed preference characterisation of exact linear aggregation à la Gorman (1953, 1961), which implies a normative representative consumer independent of the income distribution.

Individual rationality. Suppose that we have a balanced microdata panel of consumers indexed by $h = 1, ..., H$ observed over a number of periods indexed $t = 1, ..., T$.[2] Following Gorman (1953), we make the classical assumption that the law of one price holds and that prices are strictly positive K-vectors ($\mathbf{p}_t \in \mathbb{R}^K_{++}$). For each consumer h we observe non-negative quantities $\mathbf{q}^h_t \in \mathbb{R}^K_+$. We will denote these microdata by $\left\{\mathbf{p}_t, \mathbf{q}^h_t\right\}^{h\in\eta}_{t\in\tau}$, with $\eta =$

[2] Interestingly, such panel data are often considered in empirical applications of revealed preference methods of the type we consider here. For example, Cherchye, De Rock and Vermeulen (2009) study a panel of Russian consumers, and Crawford (2010) a panel of Spanish consumers.

$\{1, ..., H\}$ and $\tau = \{1, ..., T\}$ being the index sets for consumers and periods, respectively. We will use $\mathbf{Q}_t = \sum_{h \in \eta} \mathbf{q}_t^h$ to denote the aggregate demand vector in period t, so that the macrodata are $\{\mathbf{p}_t, \mathbf{Q}_t\}_{t \in \tau}$. Aggregate income is denoted by Y_t and is equal to $\mathbf{p}_t' \sum_{h \in \eta} \mathbf{q}_t^h$.

In what follows, we will assume that all the consumers are rational in the sense that observed demand results from the maximisation of a well-behaved utility function subject to an individual budget constraint. Throughout, we will assume utility functions that are nonsatiated, monotonically increasing, concave and continuous.

Definition 1 (Individual rationalisation) *A utility function u^h provides an* individual rationalisation *of the data $\{\mathbf{p}_t, \mathbf{q}_t^h\}_{t \in \tau}$ for the h'th consumer if for each observation $t \in \tau$ we have $u^h(\mathbf{q}_t^h) \geq u^h(\mathbf{q})$ for all $\mathbf{q}$ with $\mathbf{p}_t'\mathbf{q} \leq \mathbf{p}_t'\mathbf{q}_t^h$.*

Before we focus on aggregate demand, it is useful to discuss the empirical content of individual rationalisation. A core result in the revealed preference approach to demand is that there exists a utility function that provides an individual rationalisation of the data $\{\mathbf{p}_t, \mathbf{q}_t^h\}_{t \in \tau}$ if and only if the data satisfy the *Generalised Axiom of Revealed Preference* (GARP).

Definition 2 (GARP) *The data $\{\mathbf{p}_t, \mathbf{q}_t^h\}_{t \in \tau}$ satisfy GARP if there exist relations R_0^h, R^h that meet:*
(A) if $\mathbf{p}_s'\mathbf{q}_s^h \geq \mathbf{p}_s'\mathbf{q}_t^h$ then $\mathbf{q}_s^h \ R_0^h \ \mathbf{q}_t^h$;
(B) if $\mathbf{q}_s^h \ R_0^h \ \mathbf{q}_u^h$, $\mathbf{q}_u^h \ R_0^h \ \mathbf{q}_v^h$, ..., $\mathbf{q}_z^h \ R_0^h \ \mathbf{q}_t^h$ for some (possibly empty) sequence $(u, v, ..., z)$ then $\mathbf{q}_s^h \ R^h \ \mathbf{q}_t^h$;
(C) if $\mathbf{q}_s^h \ R^h \ \mathbf{q}_t^h$ then $\mathbf{p}_t'\mathbf{q}_t^h \leq \mathbf{p}_t'\mathbf{q}_s^h$.

In other words, the bundle of quantities $\mathbf{q}_s^h$ is directly revealed preferred over the bundle $\mathbf{q}_t^h$ (i.e. $\mathbf{q}_s^h \ R_0^h \ \mathbf{q}_t^h$) if $\mathbf{q}_s^h$ were chosen when $\mathbf{q}_t^h$ were equally attainable (i.e. $\mathbf{p}_s'\mathbf{q}_s^h \geq \mathbf{p}_s'\mathbf{q}_t^h$); see condition (A). Next, the *revealed preference* relation R^h exploits transitivity of preferences; see condition (B). Finally, condition (C) imposes that the bundle of quantities $\mathbf{q}_t^h$ cannot be more expensive than revealed preferred quantities $\mathbf{q}_s^h$.

We can now state the following result, which is usually referred to as Afriat's Theorem (Varian, 1982; based on Afriat, 1967):

Theorem 1 (Afriat's Theorem) *The following statements are equivalent:*
(1.A). There exists a nonsatiated, monotonic, concave and continuous utility function u^h that provides an individual rationalisation of the data $\{\mathbf{p}_t, \mathbf{q}_t^h\}_{t \in \tau}$.
(1.B) The data $\{\mathbf{p}_t, \mathbf{q}_t^h\}_{t \in \tau}$ satisfy GARP.
(1.C) For all $s, t \in \tau$, there exist numbers u_s^h, $u_t^h \in \mathbb{R}_+$ and $\beta_t^h \in \mathbb{R}_{++}$ that meet the Afriat inequalities

$$u_s^h \leq u_t^t + \beta_t^h \mathbf{p}_t' (\mathbf{q}_s^h - \mathbf{q}_t^h).$$

The equivalence between statements $(1.A)$ and $(1.B)$ captures what we mentioned above: any data set $\{\mathbf{p}_t, \mathbf{q}_t^h\}_{t \in \tau}$ can be rationalised by a well-behaved utility function if and only if these price-quantity pairs satisfy GARP. Next, the equivalent statement $(1.C)$ defines so-called Afriat inequalities, which are expressed in the unknowns u_t^h and β_t^h. These Afriat inequalities allow us to obtain an explicit construction of the utility levels and the marginal utility of income associated with each observation t: they define a utility level u_t^h and a marginal utility of income β_t^h (associated with the observed income $\mathbf{p}_t'\mathbf{q}_t^h$) for each observed $\mathbf{q}_t^h$. Importantly, as has been demonstrated by Varian (1982), and later by Blundell, Browning and Crawford (2003, 2008) and Blundell et al. (2015), the above insights can be used to formally evaluate policy reforms in terms of individual welfare.

Let us then consider rationalising the data ($\{\mathbf{p}_t, \mathbf{q}_t^h\}_{t\in\mathcal{T}}^{h\in\eta}$ and $\{\mathbf{p}_t, \mathbf{Q}_t\}_{t\in\mathcal{T}}$) in terms of a representative consumer. An important thing to note here is that there are actually two main personifications of this representative consumer.

The positive representative consumer. The *positive* representative consumer exists whenever aggregate demand can be modelled as the outcome of rational, maximising behaviour given prices and aggregate income. The positive representative consumer can be thought of as having classically well-behaved preferences, but those preferences need not have any normative significance.[3] As Gorman (1976) aptly put it, the positive representative consumer is "rather an odd chap ...he is as likely as not to be radiantly happy when those he represents are miserable and *vice versa*".[4]

The revealed preference characterisation of this "odd chap" was given by Varian (1984) and turned out to be simple: the macrodata $\{\mathbf{p}_t, \mathbf{Q}_t\}_{t\in\mathcal{T}}$ must satisfy GARP. This is very easily testable and does not involve any parametric assumptions about the form of the macro-utility function.[5]

Whilst the positive representative consumer is a potentially useful entity upon which one can base macro-level predictions, the trouble with him is, as Gorman (1976) was pointing out, that he is not fully "representative" in the welfare sense - none of the implied aggregate utility functions associated with his preferences can necessarily be thought of as a social welfare function. As a result the positive version of the representative consumer cannot be used for welfare analysis.

The normative representative consumer. The normative representative consumer is a special case of the positive representative consumer. Like the positive consumer he also exists whenever aggregate demands can be modelled as the outcome of rational, maximising behaviour given prices and aggregate income. However, the normative consumer's preferences can properly be regarded as an aggregate social welfare function. This makes him a much more useful construction: you can use him both to make predictions and to make welfare statements. The normative representative consumer is modelled as solving the following problem:[6]

$$\max_{\mathbf{q}^1,\ldots,\mathbf{q}^H\in\mathbb{R}_+^K} V\left(u^1\left(\mathbf{q}^1\right),\ldots,u^H\left(\mathbf{q}^H\right)\right) \text{ subject to } \mathbf{p}_t'\sum_{h=1}^{H}\mathbf{q}^h = Y_t, \tag{1}$$

where Y_t is aggregate income and where $u^1,\ldots,u^H$ and V are well-behaved utility functions. The question we focus on concerns the conditions under which the microdata and the asso-

[3] See, for example, Dow and Werlang (1988), Kirman (1992) and Jerison (1994).

[4] See Gorman (1976), reprinted in Blackorby and Shorrocks (1995).

[5] See, for example, Crawford and Neary (2008) for an application to country level consumption data.

[6] See, for example, Mas-Colell, Whinston and Green (1995), 4.D.1[B], p.125. We note that the normative representative consumer's utility function has the same structure as a latently separable (Gorman, 1968, 1978, Blundell and Robin, 2000, and Crawford, 2006) utility function - except for the important difference that the micro-level allocations to individuals are not latent; they are observed.

ciated macro behaviour can be rationalised by this model. In what follows, we derive these conditions under the assumption that some income distribution rule guarantees a socially optimal distribution of the aggregate income over the individual consumers. We return to this income distribution rule concept in more detail at the end of this section.

The following defines what it means for data to be rationalised by the preferences of a normative representative consumer (when assuming a socially optimal income distribution rule).

Definition 3 (Normative representative consumer rationalisation) *The utility functions V, u^1, ... , u^H provide a normative representative consumer rationalisation of the data $\left\{\mathbf{p}_t, \mathbf{q}_t^h\right\}_{t\in\tau}^{h\in\eta}$ if $V(u^1(\mathbf{q}_t^1), \ldots, u^H(\mathbf{q}_t^H)) \geq V(u^1(\mathbf{q}^1), \ldots, u^H(\mathbf{q}^H))$ for all alternative micro-allocations $\left\{\mathbf{q}^h\right\}^{h\in\eta}$ such that $\mathbf{p}_t' \sum_{h=1}^H \mathbf{q}_t^h \geq \mathbf{p}_t' \sum_{h=1}^H \mathbf{q}^h$.*

This is simply a statement of the principle of revealed preference in the relevant context: that the normative representative consumer's utility function should associate a higher real number with the observed allocation of resources than it does for any affordable alternative allocation. The next result presents the conditions under which there exists a normative representative consumer who rationalises the data (the proofs of this and all of the following results are in the Appendix).

Theorem 2 *The following statements are equivalent:*
(2.A) There exist nonsatiated, monotonic, concave and continuous utility functions $V, u^1, ..., u^H$ that provide a normative representative consumer rationalisation of the data $\left\{\mathbf{p}_t, \mathbf{q}_t^h\right\}_{t\in\tau}^{h\in\eta}$.
(2.B) For all $s, t \in \tau$ and $h \in \eta$, there exist numbers $V_s, V_t, u_s^h, u_t^h \in \mathbb{R}_+$ and $\mu_t, b_t^h \in \mathbb{R}_{++}$ such that

$$V_s \leq V_t + \mu_t \mathbf{b}_t' (\mathbf{u}_s - \mathbf{u}_t), \tag{2.B.1}$$

$$u_s^h \leq u_t^h + \frac{1}{b_t^h} \mathbf{p}_t' \left(\mathbf{q}_s^h - \mathbf{q}_t^h\right), \tag{2.B.2}$$

with $\mathbf{u}_t = (u_t^1, \ldots, u_t^H)'$ and $\mathbf{b}_t = (b_t^1, \ldots, b_t^H)'$.

Some remarks are in order. Firstly, similar to before, this is an equivalence result, so the conditions in statement $(2.B)$ are both necessary and sufficient: if there exist solutions to the inequalities then the microdata are exactly reproducible by the model of the normative representative consumer with suitable, well-behaved utility functions; equally, if solutions to these inequalities do not exist then neither do suitable, well-behaved utility functions capable of rationalising the data. Secondly, condition $(2.B.2)$ is an Afriat inequality which applies to each consumer in the microdata, and it is equivalent to the statement that the microdata on each consumer, taken one-at-a-time, satisfies GARP. What this means is that it is a necessary condition that every consumer is rationalisable by a well-behaved, individual utility function. This, of course, is entirely natural: if the representative consumer is to be normatively significant, it is clearly necessary that those he is intended to represent are themselves rationalisable. Note that individual preferences are allowed to be arbitrarily heterogeneous across consumers and can take any form - the only restrictions are that these individual preferences must be rational and well-behaved. Thirdly, condition $(2.B.1)$ is an Afriat inequality that captures the existence of a well-behaved utility function that aggregates the consumer's utility functions. Finally, whilst the form of Theorem 2 is entirely different to the kind of results found in the exact aggregation literature, which makes use of shape restrictions (there are no functional forms, and there is nothing which relates in an obvious way to the marginal utility of income), the Afriat numbers in statement $(2.B)$ bear certain important interpretations which do relate to the standard approach. The numbers $\left\{u_t^h, 1/b_t^h\right\}_{t\in\tau}$, for example, can be interpreted as utility levels and the marginal utility of income at each observed choice for consumer h.[7] Similarly, the numbers $\{V_t, \mu_t\}_{t\in\tau}$ can be interpreted as a measure of aggregate welfare and the marginal social utility of income. Note that neither the distribution of the marginal utility of individual income or the marginal social utility of income are restricted other than via their interaction in $(2.B.1)$. This interaction is important, however, so we turn to it next.

The conditions in $(2.B)$ provide a characterisation of the necessary and sufficient empirical conditions for a normative representative consumer. They are also very general - there are no restrictions on micro-preferences other than well-behavedness and none at all on the type or distribution of unobservable heterogeneity. However, there is a difficulty: these conditions are not fully testable. This is because the Afriat numbers in $(2.B)$ are not unique. What this means in practice is that as soon as the investigator finds a solution to the inequalities, the search stops and a normative representative consumer is known to exist. However, if after searching for a while no solution has been found, the only option is to keep searching. Unfortunately, the set of possible Afriat numbers is infinite and it would take forever to exhaust them. Conditions like this are sometimes said to have a bias towards acceptance - simply because a falsification result would take an infinite amount of time to determine while an acceptance, by definition, does not.[8]

The difficulty can be thought of as follows: in order for the observed distribution of resources to be optimal, the representative consumer needs to equalise the marginal social utility of income across consumers. Arguing loosely from the chain rule, marginal social utility can be thought of as the individual's marginal utility of income multiplied by the marginal contribution of individual utility to social utility (i.e. $\mu_t = \left(1/b_t^h\right) \nabla V\left(u_t^h\right)$). Therefore the term $\mu_t \mathbf{b}_t$ represents a tangle of unobservables which make $(2.B.1)$ nonlinear in unknowns. It is this which gives rise to the problem of infinite testability.

In order to make progress towards a computationally feasible necessary and sufficient condition it is going to be necessary to simplify the interaction between individual marginal utility and social weights. We explore this further next.

[7] To explain more in detail: given that the individual utility function u^h is concave (and assuming differentiability for ease of exposition, though this is easily relaxed), we have that $u^h\left(\mathbf{q}_s^h\right) \leq u^h\left(\mathbf{q}_t^h\right) + \boldsymbol{\nabla} u^h\left(\mathbf{q}_t^h\right)'\left(\mathbf{q}_s^h - \mathbf{q}_t^h\right)$ for all s, t. Maximising behaviour implies that the usual first order conditions are $\boldsymbol{\nabla} u^h\left(\mathbf{q}_t^h\right) \leq \left(1/b_t^h\right)\mathbf{p}_t$ (allowing for non-purchase of some goods), where $1/b_t^h$ represents the value of the Lagrange multiplier in the budget constraint. We can substitute this into the concavity condition to give $u^h\left(\mathbf{q}_s^h\right) \leq u^h\left(\mathbf{q}_t^h\right) + \left(1/b_t^h\right)\mathbf{p}_t'\left(\mathbf{q}_s^h - \mathbf{q}_t^h\right)$. This has the same form as condition $(2.B.2)$. So maximisation of the real-valued utility function means that there exist real numbers $u_t^h = u^h\left(\mathbf{q}_t^h\right)$ and $1/b_t^h$ which bear the required interpretation. See Varian (1982) for further discussion.

[8] This problem is closely related to revealed preference tests for weak separability (Varian, 1983). Also these necessary and sufficient tests turn out to be based on a nonlinear system of inequalities, which is empirically less attractive. See Cherchye et al. (2015) and Echenique (2015) for formal statements about the computational complexity of weak separability. A number of alternative separability tests have been proposed, which are either necessary or sufficient. See, for example, Swofford and Whitney (1987, 1994) and Fleissig and Whitney (2003, 2008).

Theorem 3 *The following statements are equivalent:*
(3.A) There exist nonsatiated, monotonic, concave and continuous utility functions V, $u^1, ..., u^H$, with common marginal utility of income, that provide a normative representative consumer rationalisation of the data $\{\mathbf{p}_t, \mathbf{q}_t^h\}_{t\in\tau}^{h\in\eta}$.
(3.B) There exist nonsatiated, monotonic, concave and continuous utility functions V, $u^1, ..., u^H$, for which V is additively separable in $u^1, ..., u^H$, that provide a normative representative consumer rationalisation of the data $\{\mathbf{p}_t, \mathbf{q}_t^h\}_{t\in\tau}^{h\in\eta}$.
(3.C) For all $s, t \in \tau$ and $h \in \eta$, there exist numbers $u_s^h, u_t^h \in \mathbb{R}_+$ and $b_t \in \mathbb{R}_{++}$ such that

$$u_s^h \leq u_t^h + \frac{1}{b_t}\mathbf{p}_t'\left(\mathbf{q}_s^h - \mathbf{q}_t^h\right).$$

What this result says is that we can either tie down the social weights to be the same across consumers (i.e. have a utilitarian social welfare function; statement $(3.B)$) or we can tie down the marginal utility of income to be the same across consumers (statement $(3.A)$). Either way, what this does is simplify the inequalities in Theorem 2 to a single (and crucially) *linear* problem (statement $(3.C)$). This inequality is very straightforward to test and does not suffer from the bias-towards-acceptance problem - it is determinable in finite time.

Income distribution rule. To conclude this section, it is important to emphasise that Theorems 2 and 3 both imply the existence of an income distribution rule that distributes aggregate income optimally from a social point of view (i.e. in the sense of Samuelson, 1956, and according to the social welfare function in (1)). Formally (and within the framework of the shape restriction based literature), an income distribution rule is a family of functions $\left(w^1(\mathbf{p},Y), w^2(\mathbf{p},Y), ..., w^H(\mathbf{p},Y)\right)$ such that $\sum_{h\in\eta} w^h(\mathbf{p},Y) = Y$ for all $\mathbf{p}$ and Y. In case there is an income distribution rule, then aggregate demand can always (and trivially) be written as a function of aggregate income through $\mathbf{Q} = \sum_{h\in\eta} \mathbf{g}^h\left(\mathbf{p}, w^h(\mathbf{p},Y)\right)$, where $\mathbf{g}^h(.,.)$ is consumer h's vector-valued demand associated with this consumer's preferences. Further, aggregate demand is the result of the representative consumer's preference relation that is represented by the social welfare function (1). Consequently, Theorems 2 and 3 imply constraints on the possible income distributions in general; this is because the aggregate demand generally depends on the income distribution rule (see Samuelson, 1956, Jerison, 1994, and Mas-Colell, Whinston and Green, 1995, for further discussion). In the next section, we consider the same question but now we will consider the existence of a normative representative consumer independent of the income distribution. This is essentially the question that Gorman (1953) originally addressed: it asks for the revealed preference conditions associated with exact linear aggregation. Interestingly, we will show that the conditions in Theorem 3 also characterise the Gorman-type normative representative consumer under a very weak data requirement.

3 Exact linear aggregation

We next investigate the conditions needed to guarantee exact linear aggregation, i.e. aggregate demand only depends on aggregate income and is not affected by how the income is actually distributed across consumers. From the shape restriction-based literature, we know that this independence result applies *if and only if* consumers have preferences of the Gorman Polar Form and linear Engel curves with common slopes. As demonstrated by Gorman (1953, 1961), this implies that aggregate demand can be written in the simple form $\mathbf{Q} = \mathbf{g}(\mathbf{p}, Y)$, where $\mathbf{g}(.,.)$ is the vector-valued demand equation that results from the maximisation of the normative representative consumer's preferences given aggregate income Y and taking as given market prices $\mathbf{p}$. Clearly, this requires that *any* income distribution, such that $\sum_{h \in \eta} y^h = Y$, gives rise to the same aggregate demands $\mathbf{Q}$.

Our following discussion will show a close link between the result in Theorem 3 and Gorman-type aggregation. We proceed in four steps. Firstly, we derive a revealed preference characterisation of individual preferences of the Gorman Polar Form. Secondly, we show the remarkable and important result that if observed prices are nonproportional, then GARP is equivalent to having preferences of the Gorman Polar Form. In practice, this data requirement is very weak: we are not aware of observational (non-experimentally generated) panel data on consumer behaviour which exhibits price-proportionality. Next, we provide the revealed preference counterpart to Gorman's aggregation results and show that, in the revealed preference sense, aggregate demand is independent of the income distribution if and only if all consumers have preferences of the Gorman Polar Form with common marginal utilities of income. In other words, all consumers are associated with parallel linear Engel curves. Finally, we propose an easy-to-apply linear test for a normative representative consumer, which holds for any possible income distribution, by combining the above steps. Interestingly, as we will discuss, the linear condition that is tested is empirically equivalent to the condition $(3.C)$ in Theorem 3.

Gorman Polar Form preferences. We begin by defining what it means for the data of an individual consumer to be rationalisable with the Gorman Polar Form. The Gorman Polar Form is usually defined in terms of an indirect utility function w^h. Let y^h represent the income of consumer h. The indirect utility function w^h is connected with the utility function u^h in the following way:

$$w^h\left(\mathbf{p}, y^h\right) = \max_{\mathbf{q}^h} \{u^h\left(\mathbf{q}^h\right) | \mathbf{p}'\mathbf{q}^h \leq y^h\}.$$

We can now state the next definition.

Definition 4 (Gorman Polar Form Rationalisation) *The data $\left\{\mathbf{p}_t, \mathbf{q}_t^h\right\}_{t \in \mathcal{T}}$ are rationalisable by the Gorman Polar Form if they are rationalisable by a utility function u^h (in the sense of Definition 1) such that the indirect utility function $w^h(\mathbf{p}, y^h) = \frac{y^h - a^h(\mathbf{p})}{b^h(\mathbf{p})}$, with $a^h(\mathbf{p}) \in \mathbb{R}$ and $b^h(\mathbf{p}) \in \mathbb{R}_{++}$ for all $\mathbf{p}$ and the functions a^h and b^h linearly homogeneous of degree 1.*

In this definition, the price index $a^h(\mathbf{p})$ is often interpreted as subsistence expenditure - although this interpretation is not always valid (see Pollak, 1971, p 403, fn 4) - while the price index $b^h(\mathbf{p})$ is interpreted as the inverse of the marginal utility of income.

Before moving on, it is worth pointing out the well-established fact that the Gorman Polar Form does not necessarily give rise to well-behaved preferences in all parts of the quantity-space: in general, well-behaved preferences only apply to a limited range of possible income values. For instance, for some income values, the linear Engel curves may lead to negative consumption or cross with each other. To avoid such problems, Gorman Polar Form preferences are usually defined subject to bounds on possible income levels.[9] To keep the exposition simple, our following analysis only considers income values that lie within such income ranges and, thus, we will not explicitly consider income bounds in our exposition (but, importantly, bounds on income levels do appear in the proof of Theorem 4). We can then state the characterisation.[10]

Theorem 4 *The following statements are equivalent:*
(4.A) The data $\{\mathbf{p}_t, \mathbf{q}_t^h\}_{t\in\tau}$ *are rationalisable by the Gorman Polar Form.*
(4.B) For all $s, t \in \tau$*, there exist numbers* w_s^h*,* $w_t^h \in \mathbb{R}_+$*,* $a_t^h \in \mathbb{R}$ *and* $b_t^h \in \mathbb{R}_{++}$ *such that*

$$w_s^h \leq w_t^h + \frac{1}{b_t^h}\mathbf{p}_t' \left(\mathbf{q}_s^h - \mathbf{q}_t^h\right), \tag{4.B.1}$$

$$w_t^h = \frac{\left(\mathbf{p}_t'\mathbf{q}_t^h\right) - a_t^h}{b_t^h}, \tag{4.B.2}$$

$$a_t^h = \delta a_s^h \textit{ and } b_t^h = \delta b_s^h \textit{ if } \mathbf{p}_t = \delta \mathbf{p}_s \textit{ for } \delta > 0. \tag{4.B.3}$$

As before the Afriat numbers in this result have certain structural interpretations. Condition $(4.B.1)$, for example, is again an Afriat inequality, which has a directly similar interpretation as before. In this inequality, we can interpret each number w_t^h as an indirect utility value (the function value $w^h(\mathbf{p}, y^h)$ in Definition 4, which equals the utility value $u^h\left(\mathbf{q}^h\right)$ under rational consumer behaviour). Condition $(4.B.2)$ then states the Gorman Polar Form restriction, with the numbers a_t^h and b_t^h corresponding to the price indices $a^h(\mathbf{p})$ and $b^h(\mathbf{p})$ in Definition 4 evaluated at $\mathbf{p}_t$. Condition $(4.B.3)$, finally, imposes linear homogeneity of these price indices.

Two further notes are in order. First, the Gorman Polar Form characterisation in Theorem 4 is nonlinear in a_t^h and b_t^h. However, in our proof of Theorem 4 we show that it can be equivalently expressed in linear form. In turn, this makes it easily testable.

The second remark combines the results in Theorems 1 and 4. In particular, it follows that, under the weak data requirement of nonproportional prices, Gorman Polar Form preferences provide *no additional restrictions* over and above the standard Afriat inequalities (or, equivalently, GARP).[11] In other words, Gorman Polar Form preferences and rational preferences are nonparametrically (in the revealed preference sense) *equivalent*: for data in which proportional prices movements are not observed their empirical implications are identical. This result is formally stated as follows:

---

[9] See, for example, Pollak (1971) and Blackorby, Boyce and Russell (1978) for a more detailed discussion of the local nature of Gorman Polar Form preferences.

[10] An alternative revealed preference characterisation of the Gorman Polar Form can be found in work in progress by Brown and Shannon. In a certain sense, the work of these authors is complementary to ours as Brown and Shannon characterise Gorman Polar Form preferences in terms of so-called 'dual' Afriat numbers (which have an interpretation in terms of indirect utility functions; see Brown and Shannon, 2000), whereas our analysis starts from the original 'primal' Afriat numbers (to be interpreted in terms of direct utility functions). We thank Don Brown for revealing this to us in a private conversation.

[11] Specifically, under nonproportional prices condition $(4.B.3)$ becomes redundant. Then, one can easily verify that, for any given solution for the Afriat inequalities $(4.B.1)$, there also exists a solution for condition $(4.B.2)$.

Corollary 1 *The following statements are equivalent when prices* $\mathbf{p}_t \neq \delta \mathbf{p}_s$ $(\delta > 0)$ *for all* $s, t \in \tau$*:*
(A) The data $\left\{\mathbf{p}_t, \mathbf{q}_t^h\right\}_{t \in \tau}$ *are rationalisable by the Gorman Polar Form.*
(B) The data $\left\{\mathbf{p}_t, \mathbf{q}_t^h\right\}_{t \in \tau}$ *satisfy GARP.*

This is an important result. It implies that if the data satisfy GARP and observed prices are nonproportional, then we can *always* construct an indirect utility function which exactly rationalises the data with the Gorman Polar Form. This is perhaps surprising as the Gorman Polar Form is usually thought of as a very demanding restriction. However, it seems that this is only the case when proportional prices are observed in the data. In such a case, the Gorman Polar Form is *extremely* demanding as we can directly observe points on an Engel curve and this Engel curve must be perfectly straight. However, as indicated above, we are not aware of *any* observational (non-experimentally generated) consumer panel data in which proportional prices changes are ever observed. Thus, it turns out that, empirically, the Gorman Polar Form is without additional empirical content from a revealed preference point of view.[12]

Exact linear aggregation. We can now use these insights to provide the revealed preference counterparts of Gorman's conditions for exact linear aggregation. As stressed above, this implies that aggregate demand is independent of the income distribution. Gorman proved that such exact aggregation holds if and only if consumers' preferences are of the Gorman Polar Form with common slopes for the (linear) Engel curves. In revealed preference terms, we get the following characterisation.

Theorem 5 *The following statements are equivalent for the data* $\left\{\mathbf{p}_t, \mathbf{q}_t^h\right\}_{t \in \tau}^{h \in \eta}$*:*
(5.A) Aggregate demand is independent of the income distribution.
(5.B) For all $s, t \in \tau$ *and* $h \in \eta$*, there exist numbers* w_s^h*,* $w_t^h \in \mathbb{R}_+$*,* $a_t^h \in \mathbb{R}$ *and* $b_t \in \mathbb{R}_{++}$ *such that*

$$w_s^h \leq w_t^h + \frac{1}{b_t} \mathbf{p}_t' \left(\mathbf{q}_s^h - \mathbf{q}_t^h\right), \tag{5.B.1}$$

$$w_t^h = \frac{\left(\mathbf{p}_t' \mathbf{q}_t^h\right) - a_t^h}{b_t}, \tag{5.B.2}$$

$$a_t^h = \delta a_s^h \text{ and } b_t = \delta b_s \text{ if } \mathbf{p}_t = \delta \mathbf{p}_s \text{ for } \delta > 0. \tag{5.B.3}$$

As compared to Theorem 4, the key requirement is that the Afriat number b_t is common across consumers who face the same prices (i.e. $b_t^h = b_t$ for all h). In terms of Definition 4, this effectively imposes Gorman Polar Form preferences with a common $b(\mathbf{p})$ index for all consumers. The idea is that the marginal utility of income must be independent of income variations across consumers but can vary with prices. Without these restrictions on the individual preferences (and, by implication, on the preferences of the normative representative consumer), one typically has to assume some income distribution rule (as discussed in Section 2). We note, finally, that our characterisation in Theorem 5 can be linearised in a directly

[12] At this point it is worth recalling that we focus on preferences taking the Gorman Polar Form for income values within bounded ranges, which here means that the equivalence in Corollary 1 has a local nature by construction.

similar way as our earlier characterisation in Theorem 4. As such, it implies an easy-to-apply test for a normative representative consumer that is independent of the income distribution.

Interestingly, the characterisation in Theorem 5 also generalises several special cases that generate the same independence of the income distribution. Two important examples are Varian's (1983) revealed preference characterisation of identical homothetic preferences (where $a^h(\mathbf{p}) = 0$ in Definition 4) and Brown and Calsamiglia's (2007) revealed preference characterisation of quasi-linear preferences (where $a^h(\mathbf{p}) = -p^i\phi(\mathbf{p})$ and $b^h(\mathbf{p}) = p^i$, with p^i the price of the numeraire and ϕ a homogeneous of degree one function).

As a final result, we connect the characterisations in Theorems 3 and 5. Similar to Corollary 1, we find that if observed prices are nonproportional, then a necessary and sufficient condition for a Gorman-type normative representative consumer is that each consumer satisfies the standard Afriat inequalities with a common marginal utility of income. This is formally stated in the following result:

Corollary 2 *The following statements are equivalent when prices $\mathbf{p}_t \neq \delta\mathbf{p}_s$ $(\delta > 0)$ for all $s, t \in \tau$:*
(A) Aggregate demand is independent of the income distribution.
(B) For all $s, t \in \tau$ and $h \in \eta$, there exist numbers w_s^h, $w_t^h \in \mathbb{R}_{++}$ and $b_t \in \mathbb{R}_{++}$ such that

$$w_s^h \leq w_t^h + \frac{1}{b_t}\mathbf{p}_t'\left(\mathbf{q}_s^h - \mathbf{q}_t^h\right).$$

Thus, we get exactly condition $(3.C)$ for aggregate demand to be independent of the income distribution. This means that, under nonproportional prices, the condition in Theorem 5 conveniently reduces to the condition in Theorem 3. In other words, under the weak data requirement of nonproportional prices, the characterisation of a normative representative consumer in Theorem 3 holds for all income distributions across consumers and no longer relies on the existence of an income distribution rule. On the other hand, if prices are proportional, then the condition in Corollary 2 (or condition $(3.C)$ in Theorem 3) is not empirically equivalent to the one in Theorem 5. In that case, it still (but only) defines a necessary (and not sufficient) test for exact linear aggregation: if the condition is violated we can (only) conclude that there certainly does not exist a normative representative consumer that is independent of the income distribution.

4 Conclusion

The concept of the normatively significant representative consumer has long played a central role in many areas in economics. Although the conditions for existence have been argued to be demanding, it is fair to say that existing evidence is mainly based on Gorman's well-known exact linear aggregation results which involve a shape restriction. To test econometrically Gorman's conditions for exact linear aggregation (which boil down to consumers having preferences of the Gorman Polar Form with an equal marginal utility of income), one needs to make many additional assumptions to bring these conditions to the data.

In this paper, we revisited the exact aggregation problem by bringing in tools from the revealed preference literature. These tools are based solely on the data at hand and do not need any additional parametric or statistical assumptions. As such, they will allow empirical researchers to robustly analyse the empirical validity of exact aggregation.

Our main theoretical contribution is that we provide a number of closely linked results relating to the existence of a consumer that can normatively represent a group of consumers, regardless of the income distribution. Usefully, the most important conditions are linear and thus easy to apply in practice. Our analysis also clarified the relationship between the empirical restrictions associated with Samuelson-type aggregation and Gorman-type aggregation. Most notably, we made explicit the empirical conditions under which the two notions of aggregation become equivalent.

Appendix

Proof of Theorem 2.

$(2.A) \Rightarrow (2.B)$: First consider the implications of optimising behaviour and the first order conditions from the consumer's problem. Continuity ensures that suitable subgradients exist such that $\boldsymbol{\nabla} V\left(\mathbf{q}_t^h\right) \leq \mu_t \mathbf{p}_t$ where $\boldsymbol{\nabla} V\left(\mathbf{q}_t^h\right) = \nabla V\left(u_t^h\right) \boldsymbol{\nabla} u^h\left(\mathbf{q}_t^h\right)$. Define $\mu_t b_t^h = \nabla V\left(u_t^h\right)$. Then $\boldsymbol{\nabla} u^h\left(\mathbf{q}_t^h\right) \leq \left(b_t^h\right)^{-1} \mathbf{p}_t$. Now consider the concavity conditions for this structure

$$\begin{aligned} V\left(\mathbf{u}_s\right) &\leq V\left(\mathbf{u}_t\right) + \boldsymbol{\nabla} V\left(\mathbf{u}_t\right)' \left(\mathbf{u}_s - \mathbf{u}_t\right) \\ u^h\left(\mathbf{q}_s^h\right) &\leq u^h\left(\mathbf{q}_t^h\right) + \boldsymbol{\nabla} u^h\left(\mathbf{q}_t^h\right)' \left(\mathbf{q}_s^h - \mathbf{q}_t^h\right) \end{aligned}$$

Substituting in $\boldsymbol{\nabla} u^h\left(\mathbf{q}_t^h\right) \leq \left(b_t^h\right)^{-1} \mathbf{p}_t$ and $\mu_t b_t^h = \nabla V\left(u_t^h\right)$ preserves the inequalities and gives

$$\begin{aligned} V\left(\mathbf{u}_s\right) &\leq V\left(\mathbf{u}_t\right) + \mu_t \mathbf{b}_t' \left(\mathbf{u}_s - \mathbf{u}_t\right) \\ u^h\left(\mathbf{q}_s^h\right) &\leq u^h\left(\mathbf{q}_t^h\right) + \frac{1}{b_t^h} \mathbf{p}_t' \left(\mathbf{q}_s^h - \mathbf{q}_t^h\right) \end{aligned}$$

which are conditions $(2.B.1)$ and $(2.B.2)$.

$(2.B) \Rightarrow (2.A)$: Suppose we have numbers $\{V_t, \mu_t > 0\}_{t \in \tau}$ and $H-$vectors $\{\mathbf{u}_t, \mathbf{b}_t > 0\}_{t \in \tau}$ such that conditions $(2.B.1)$ and $(2.B.2)$ hold. Consider some arbitrary $\left\{\mathbf{q}^h\right\}^{h \in \eta}$ such that $\mathbf{p}_t' \sum \mathbf{q}_t^h \geq \mathbf{p}_t' \sum_h \mathbf{q}^h$ for some observation t. We need to show that there exists utility functions, with the stated properties such that $V\left(u^1\left(\mathbf{q}_t^1\right), ..., u^H\left(\mathbf{q}_t^H\right)\right) \geq V\left(u^1\left(\mathbf{q}^1\right), ..., u^H\left(\mathbf{q}^H\right)\right)$. Using $(2.B.2)$ we can construct T upper bounds on $u^h\left(\mathbf{q}^h\right)$ and if we take the minimum of these then we have, as in Varian (1982), a piecewise linear, nonsatiated, monotonic, concave and continuous utility function

$$u^h\left(\mathbf{q}^h\right) = \min_s \left\{u_s^h + \frac{1}{b_s^h} \mathbf{p}_s' \left(\mathbf{q}^h - \mathbf{q}_s^h\right)\right\}_{s \in \tau} \leq u_t^h + \frac{1}{b_t^h} \mathbf{p}_t' \left(\mathbf{q}^h - \mathbf{q}_t^h\right).$$

Summing this inequality over h after multiplying it with the strict positive number b_t^h gives

$$\mathbf{b}_t' \mathbf{u}_t - \mathbf{p}_t' \sum \mathbf{q}_t^h \geq \mathbf{b}_t' \mathbf{u} - \mathbf{p}_t' \sum \mathbf{q}^h$$

where $\mathbf{u}_t = \left(u_t^1, ..., u_t^H\right)'$, $\mathbf{u} = \left(u^1, ..., u^H\right)'$, $u^h = u^h\left(\mathbf{q}^h\right)$ and $\mathbf{b}_t = \left(b_t^1, ...b_t^H\right)'$. Since $\mathbf{p}_t' \sum \mathbf{q}_t^h \geq \mathbf{p}_t' \sum \mathbf{q}^h$ we must have that $\mathbf{b}_t'\mathbf{u}_t \geq \mathbf{b}_t'\mathbf{u}$. Using $(2.B.1)$ we can then similarly construct the following macro-utility function

$$V(\mathbf{u}) = \min_s \left\{V_s + \mu_s \mathbf{b}_s' (\mathbf{u} - \mathbf{u}_s)\right\}_{s \in \mathcal{T}} \leq V_t + \mu_t \mathbf{b}_t' (\mathbf{u} - \mathbf{u}_t).$$

Since $\mu_t \mathbf{b}_t' (\mathbf{u} - \mathbf{u}_t) \leq 0$ we obtain $V(\mathbf{u}) \leq V_t$ as required.■

Proof of Theorem 3.

$(3.A) \Leftrightarrow (3.C)$. The condition $(3.C)$ is simply $(2.B.2)$ from Theorem 1 with the common marginal utility of income requirement added. Condition $(2.B.1)$ is redundant according to the following argument. Sum $(3.C)$ over h

$$\sum_h u_s^h \leq \sum_h u_t^h + \frac{1}{b_t} \sum_h \mathbf{p}_t' \left(\mathbf{q}_s^h - \mathbf{q}_t^h\right)$$

Define $V_t = \sum u_t^h$ and $\mu_t = \frac{1}{b_t}$ then

$$V_s - V_t = \mu_t b_t \left(\mathbf{1}'\mathbf{u}_s - \mathbf{1}'\mathbf{u}_t\right)$$

since $\mu_t b_t = 1$. Hence there exist numbers such that

$$V_s \leq V_t + \mu_t b_t \left(\mathbf{1}'\mathbf{u}_s - \mathbf{1}'\mathbf{u}_t\right)$$

which is $(2.B.1)$ when $b_t^h = b_t$. Thus the conditions are equivalent to those in Theorem 2 with the extra restriction that $b_t^h = b_t$.

$(3.B) \Leftrightarrow (3.C)$ Analogous to the proof of Theorem 2. However given the additive separability of V we have $\nabla V\left(u_t^i\right) = \nabla V\left(u_t^j\right)$, i.e. this derivative is constant for all i, j. So define $\mu_t b_t = \nabla V\left(u_t^h\right)$ and note the lack of the h superscript on b_t. The rest of the proof follows that for Theorem 2 to give condition $(2.B.2)$. Summing $(2.B.2)$ across h and defining $V_t = \mathbf{1}'\mathbf{u}_t$ gives

$$V_s \leq V_t + \frac{1}{b_t} \mathbf{p}_t' \left(\sum \mathbf{q}_s^h - \sum \mathbf{q}_t^h\right)$$

$$V_s = V_t + \mathbf{1}' (\mathbf{u}_s - \mathbf{u}_t)$$

which satisfies condition (2.B.1) where we interpret $\mu_t \mathbf{b}_t = \mathbf{1}$.■

Proof of Theorem 4.

As a preliminary step, we provide an equivalent linear formulation of the conditions in $(4.B)$. Let $\alpha_t = -a_t^h / b_t^h$ and $\beta_t^h = 1/b_t^h$. Then we get the following linear reformulations of the conditions $(4.B.1) - (4.B.3)$:

$$w_s^h \leq w_t^h + \beta_t^h \mathbf{p}_t' \left(\mathbf{q}_s^h - \mathbf{q}_t^h\right), \tag{4.B.1'}$$

$$w_t^h = \alpha_t + \beta_t^h \left(\mathbf{p}_t' \mathbf{q}_t^h\right), \tag{4.B.2'}$$

$$\alpha_t^h = \alpha_s^h \text{ and } \beta_t^h = \beta_s^h / \delta \text{ if } \mathbf{p}_t = \delta \mathbf{p}_s \text{ for } \delta > 0. \tag{4.B.3'}$$

.

$(4.A) \Rightarrow (4.B)$: Condition (4.B.1') readily follows Theorem 1 for a utility function u^h that rationalises the data $\{\mathbf{p}_t, \mathbf{q}_t^h\}_{t\in\tau}$. Then, we can use $w_t^h = \max_{\mathbf{q}}\{u^h(\mathbf{q})\,|\mathbf{p}_t'\mathbf{q} \leq \mathbf{p}_t'\mathbf{q}_t^h\}$ (using $\mathbf{p}_t'\mathbf{q}_t^h = y_t^h$). Given this, Definition 4 directly implies (4.B.2') and (4.B.3') when using $\alpha_t = -a^h(\mathbf{p}_t)/b^h(\mathbf{p}_t)$ and $\beta_t^h = 1/b^h(\mathbf{p}_t)$.

$(4.B) \Rightarrow (4.A)$: Consider

$$u^h(\mathbf{q}) = \min_t \; [w_t^h + \beta_t^h \mathbf{p}_t' \left(\mathbf{q} - \mathbf{q}_t^h\right)].$$

Varian (1982) has shown that this utility function rationalises the data $\{\mathbf{p}_t, \mathbf{q}_t^h\}_{t\in\tau}$. Using (4.B.2'), we have

$$u^h(\mathbf{q}) = \min_t[\alpha_t^h + \beta_t^h \mathbf{p}_t'\mathbf{q}]. \tag{2}$$

Let us then verify whether the function u^h meets Definition 4. Consider some arbitrary prices $\mathbf{p}_0$ and income y_0^h. As a preliminary step, we recall that

$$w^h(\mathbf{p}_0, y_0^h) = \max_{\mathbf{q}}\{u^h(\mathbf{q})\,|\mathbf{p}_0'\mathbf{q} \leq y_0^h\}.$$

Thus, using (2), we get

$$w^h(\mathbf{p}_0, y_0^h) = \max_{\mathbf{q}} \left\{\min_t[\alpha_t^h + \beta_t^h \mathbf{p}_t'\mathbf{q}]|\mathbf{p}_0'\mathbf{q} \leq y_0^h\right\}.$$

Dropping the min operator, we can equivalently state

$$w^h(\mathbf{p}_0, y_0^h) = \max_{w,\mathbf{q}} \left\{w|w \leq \alpha_t^h + \beta_t^h \mathbf{p}_t'\mathbf{q} \;\; (t\in\tau)\,,\, \mathbf{p}_0'\mathbf{q} \leq y_0^h\right\},$$

which obtains the linear program

$$\begin{gathered} w^h(\mathbf{p}_0, y_0^h) = \max_{w\in\mathbb{R},\mathbf{q}\in\mathbb{R}_+^N} w \\ \text{s.t.} \\ w - \beta_t^h \mathbf{p}_t'\mathbf{q} \leq \alpha_t^h \;\; (t\in\tau)\,, \\ \mathbf{p}_0'\mathbf{q} \leq y_0^h. \end{gathered} \tag{3}$$

The dual linear program is given as

$$\begin{gathered} w^h(\mathbf{p}_0, y_0^h) = \min_{\theta_t\in\mathbb{R}_+ \, (t\in\tau), \lambda\in\mathbb{R}_+} \sum\nolimits_{t=1}^T \alpha_t^h \theta_t + \lambda y_0^h \\ \text{s.t.} \\ \sum\nolimits_{t=1}^T \theta_t = 1, \\ -\sum\nolimits_{t=1}^T \theta_t \beta_t^h \mathbf{p}_t + \lambda \mathbf{p}_0 \geq 0. \end{gathered} \tag{4}$$

Let $\theta_t^* \;\; (t\in\tau)$ and λ^* define the optimum of program (4). In general, these optimal values are independent of y_0^h when y_0^h respects boundary conditions that limit the domain of y_0^h. In practice, the boundary values for y_0^h can be determined by standard methodology for sensitivity analysis of linear programming. (Technically, these bounds will correspond to the range of y_0^h (as the objective coefficient of λ) for which the optimal basic feasible solution of the linear program (4) remains constant.) These boundary conditions parallel the usual conditions that apply to indirect utility functions representing Gorman Polar Form preferences; see our discussion following Definition 4 in the main text.

Thus, because the solution of the problem (4) is independent of y_0^h (under the stated boundary conditions), we conclude that the function w^h in (4) meets the requirement in Definition 4 for

$$\lambda^* = 1/b^h(\mathbf{p}_0) \text{ and } -a^h(\mathbf{p}_0)/b^h(\mathbf{p}_0) = \sum_t \theta_t^* \alpha_t.$$

Specifically, for w^* the optimal value of linear program (4) (or, equivalently, (3)), we get

$$\begin{aligned} w^h(\mathbf{p}_0, y_0^h) &= w^* \\ &= \lambda^* y_0^h + \sum_{t=1}^{T} \theta_t^* \alpha_t^h \\ &= \frac{y_0^h - a^h(\mathbf{p}_0)}{b^h(\mathbf{p}_0)}. \end{aligned}$$

Inspection of problems (3) and (4) reveals that the price indices a^h and b^h are linearly homogeneous of degree 1 (if again the same income boundary conditions hold).■

Proof of Corollary 1.

As a first step, we note that the conditions $(4.B.2)$ and $(4.B.3)$ in Theorem 4 are void if $\mathbf{p}_t \neq \delta\mathbf{p}_s$ $(\delta > 0)$ for all s, t. As such, rationalisability by Gorman Polar Form only requires consistency with the condition $(4.B.1)$. The equivalence between the statements (A) and (B) in Corollary **1** then follows directly from the equivalence between statements $(1.B)$ and $(1.C)$ in Theorem 1.■

Proof of Theorem 5.

This follows from Theorem 4 (i.e. each household is rationalisable by the Gorman Polar Form) and the result of Gorman (i.e. the marginal utility of income is household independent, which is captured by the common b_t (i.e. $b_t^h = b_t$ for all h)). ■

Proof of Corollary 2.

The result follows from combining Corollary 1 with Theorem 5.■

References

[1] Afriat, S. (1967), "The construction of utility functions from expenditure data", *International Economic Review*, 8, 67-77.

[2] Antonelli, G. B. (1886), *Sulla Teoria Matematica dell'Economia Politica*, Pisa. English translation in J. Chipman, L. Hurwicz, M. Richter and H. Sonnenschein (eds.) (1971), *Preferences, Utility and Demand*, Harcourt Brace, New York.

[3] Blackorby, C., R. Boyce and R. Russell (1978), "Estimation of demand systems generated by the Gorman polar form: a generalization of the S-branch utility tree", *Econometrica*, 46, 345-363.

[4] Blackorby, C. and A. Shorrocks (1995), *Separability and Aggregation, Vol.1, Collected works of W.M. Gorman*, Clarendon Press, Oxford.

[5] Blundell, R., M. Browning, L. Cherchye, I. Crawford, B. De Rock and F. Vermeulen (2015), "Sharp for SARP: nonparametric bounds on counterfactual demands", *American Economic Journal: Microeconomics*, 7, 43-60.

[6] Blundell, R., M. Browning and I. Crawford (2003), "Nonparametric Engel curves and revealed preference", *Econometrica*, 71, 205-240.

[7] Blundell, R., M. Browning and I. Crawford (2008), "Best nonparametric bounds on demand responses", *Econometrica*, 76, 1227-1262.

[8] Blundell, R. and J.-M. Robin (2000), "Latent separability: grouping goods without weak separability", *Econometrica*, 68, 53-84.

[9] Brown, D. and C. Calsamiglia (2007), "The nonparametric approach to applied welfare analysis", *Economic Theory*, 31, 183-188.

[10] Brown, D. and C. Shannon (2000), "Uniqueness, stability, and comparative statics in rationalizable Walrasian markets", *Econometrica*, 68, 1529-1540.

[11] Carroll, C. (2000), "Requiem for the representative consumer? Aggregate implications of microeconomic consumption behavior", *American Economic Review*, 90, 10-115.

[12] Cherchye, L., T. Demuynck, B. De Rock, and P. Hjertstrand (2015), "Revealed preference tests for weak separability: an integer programming approach", *Journal of Econometrics*, forthcoming.

[13] Cherchye, L., B. De Rock, and F. Vermeulen (2009), "Opening the black box of intrahousehold decision-making: theory and non-parametric empirical tests of general collective consumption models", *Journal of Political Economy*, 117, 1074-1104.

[14] Chiappori, P.-A. and I. Ekeland (2011), "New developments in aggregation economics", *Annual Review of Economics*, 3, 631-668.

[15] Crawford, I. (2006), "Necessary and sufficient conditions for latent separability", University of Surrey, Department of Economics Discussion Papers 07-06.

[16] Crawford, I. and P. Neary (2008), "Testing for a reference consumer in international comparisons of living standards", *American Economic Review*, 98, 1731-1732.

[17] Crawford, I. (2010), "Habits Revealed", *Review of Economic Studies*, 77, 1382-1402

[18] Diewert, E. (1973), "Afriat and revealed preference theory", *Review of Economic Studies*, 40, 419-425.

[19] Dow, J. and S. Werlang (1988), "The consistency of welfare judgments with a representative consumer", *Journal of Economic Theory*, 44, 265-280.

[20] Echenique F. (2015), "Testing for separability is hard", CalTech working paper.

[21] Fleissig, A. and G. Whitney (2003), "A new PC-based test for Varian's weak separability conditions", *Journal of Business and Economic Statistics*, 21, 133-144.

[22] Fleissig, A. and G.A. Whitney (2008), "A nonparametric test of weak separability and consumer preferences", *Journal of Econometrics*, 147, 275 - 281.

[23] Gorman, W. (1953), "Community preference fields", *Econometrica*, 21, 63-80.

[24] Gorman, W. (1961), "On a class of preference fields", *Metroeconomica*, 13, 53-56.

[25] Gorman, W. (1968), "Measuring the quantities of fixed factors", in J.N. Wolfe. (ed.), *Value Capital and Growth: Essays in Honour of Sir John Hicks*, Edinburg University Press, Edinburgh.

[26] Gorman, W. (1976), "Muellbauer's representative consumer", unpublished manuscript reprinted in C. Blackorby and A. Shorrocks (1995, eds.), *Separability and Aggregation. Collected works of W.M. Gorman. Volume 1*, Clarendon Press, Oxford.

[27] Gorman, W. (1978), "More measures of fixed factors" in C. Blackorby and A. Shorrocks (1995, eds.), *Separability and Aggregation. Collected works of W.M. Gorman. Volume 1*, Clarendon Press, Oxford.

[28] Jerison, M. (1994), "Optimal income distribution rules and representative consumers", *Review of Economic Studies*, 61, 739-771.

[29] Kirman, A. (1992), "Whom or what does the representative individual represent?", *Journal of Economic Perspectives*, 6, 117-36.

[30] Mas-Colell, A., M. Whinston and J. Green (1995), *Microeconomic Theory*, Oxford University Press, Oxford.

[31] Muellbauer, J. (1975), "Aggregation, income distribution and consumer demand", *Review of Economic Studies*, 62, 525-543.

[32] Muellbauer, J. (1976), "Community preferences and the representative consumer", *Econometrica*, 44, 979-999.

[33] Pollak, R. (1971), "Additive utility functions and linear Engel curves", *Review of Economic Studies*, 38, 401-414.

[34] Samuelson, P. (1938), "A note on the pure theory of consumer behaviour", *Economica*, 5, 61-71.

[35] Samuelson, P. (1948), "Consumption theory in terms of revealed preference", *Economica*, 15, 243-253.

[36] Samuelson, P. (1956), "Social indifference curves", *Quarterly Journal of Economics*, 70, 1-22.

[37] Swofford, J. and G. A. Whitney (1987), "Nonparametric tests of utility maximization and weak separability for consumption, leisure, and money", *Review of Economics and Statistics*, 69, 458-464.

[38] Swofford, J.L. and G.A. Whitney (1994), "A revealed preference test for weakly separable utility maximization with incomplete adjustment", *Journal of Econometrics* 60, 235-249.

[39] Varian, H. (1982), "The nonparametric approach to demand analysis", *Econometrica*, 50, 945-972.

[40] Varian, H. (1983), "Nonparametric tests of consumer behaviour", *Review of Economic Studies*, 50, 99-110.

[41] Varian, H. (1984), "Social indifference curves and aggregate demand", *Quarterly Journal of Economics*, 99, 403-414.